PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/63, C12P 19/34, C12Q 1/68, 1/70 — A1

(11) International Publication Number: **WO 99/05293**

(43) International Publication Date: 4 February 1999 (04.02.99)

(21) International Application Number: PCT/US98/14759

(22) International Filing Date: 24 July 1998 (24.07.98)

(30) Priority Data:
08/900,347 25 July 1997 (25.07.97) US

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Application
US Not furnished (CIP)
Filed on Not furnished

(71) Applicant *(for all designated States except US)*: THE MOUNT SINAI SCHOOL OF MEDICINE OF THE CITY UNIVERSITY OF NEW YORK [US/US]; One Gustave L. Levy Place, New York, NY 10029–6574 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: BANCROFT, Frank, C. [US/US]; 51 Dewey Street, Huntington, NY 11743 (US). FLISS, Makiko [JP/US]; Apartment 6B, 1245 Park Avenue, New York, NY 10128 (US).

(74) Agents: CLARK, Richard, S. et al.; Baker & Botts, LLP, 30 Rockefeller Plaza, New York, NY 10112–0228 (US).

(81) Designated States: AU, CA, JP, US, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(54) Title: PROCEDURE FOR PREPARING ENRICHED GENE EXPRESSION LIBRARIES

(57) Abstract

The present invention relates to a method for producing a gene expression libraries enriched for a particular class of genes comprising (i) a nucleic acid preparation step and (ii) a polyclonal antiserum screening step, wherein both steps define the class of genes of interest. The present invention also relates to the use of differential hybridization protocols for production of libraries, including expression libraries, enriched for a particular class of genes. The method of the invention, using selection procedures directed not only at the underlying nucleic acid but also a functional attribute of the target class of proteins, may be used to produce gene expression libraries enriched for genes encoding a desired class of proteins, including, but not limited to, secretory proteins, cell cycle associated proteins, proteins defined by a particular subcellular location, tissue specific proteins, infection–related proteins, and non–infectious disease associated proteins.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

PROCEDURE FOR PREPARING ENRICHED GENE EXPRESSION LIBRARIES

SPECIFICATION

1. INTRODUCTION

The present invention relates to an improved method for preparing gene libraries, including expression libraries, representative of a functionally defined class of genes. In a preferred, nonlimiting example, the present invention provides methods of preparing gene libraries, including expression libraries, enriched for genes encoding secretory proteins, as well as the libraries so produced.

2. BACKGROUND OF THE INVENTION

In view of the current initiative to characterize the genomes of humans and other organisms, there is a need for methods of producing gene expression libraries representative of classes of genes of interest, in that non-specific libraries are associated with numerous inefficiencies. To date, gene expression libraries have generally been created using, as the specificity determining factor, nucleic acid selected as being representative of the desired class of genes.

For example, to prepare a tissue specific gene expression library, mRNA was prepared from the tissue or cell type of interest, converted into cDNA and inserted into an appropriate expression vector. To eliminate genes shared among tissues, the mRNA may have been annealed with mRNA from another tissue, cDNA or oligonucleotides prepared therefrom and the resulting hybrids removed prior to incorporation into the expression vector. The finished expression library could then be screened with an antibody directed toward the product of a specific gene of interest. Similar methods were used to prepare libraries using nucleic acid defined by other characteristics, for example, nucleic acids present in virally infected cells, or associated

with a particular phase of the cell cycle. The problem with these techniques is that the selection techniques do not satisfactorily assure that non-specific genes are eliminated or, conversely, that class specific genes are not lost.

It would be of particular interest to produce a gene expression library enriched for secretory proteins. Most clinically significant proteins are secreted by their tissues of origin, and then exert their action at a distant location. The genes for a number of previously known secretory proteins have now been cloned and produced in a recombinant form, including, for example, growth hormone and other growth factors and cytokines, interferons, insulin, and erythropoietin. However, it is highly likely that a large number of clinically significant proteins remains to be identified.

It has long been known that the mRNAs for secretory proteins are associated, virtually exclusively, with membrane-bound polysomes contained in the "rough microsome" cellular fraction" (Blobel and Dobberstein, 1975, J. Cell Biol. 67:835-851). The specific information for this association is contained within a "signal sequence" in the N-terminus of the nascent protein, which thus tags proteins for secretion or, in some cases, membrane insertion (Sabatini et al., 1982, Cold Spring Harbor Symp. Quant. Biol. 46:807-818). However, there is apparently no consensus amino acid sequence for the signal sequence contained in nascent secretory proteins. Accordingly, it has not been feasible to identify genes encoding secretory proteins by searching DNA sequence banks for DNA sequences that may represent signal sequences.

The purpose of the present invention is to provide a method for producing class-enriched gene libraries, including expression libraries, which define the clonal members of the libraries not only by a nucleic acid selection procedure but also by antibody-based screening based on a functional attribute of the class of gene products. In preferred embodiments this method is used to create libraries, including expression libraries, enriched for secretory proteins.

3. SUMMARY OF THE INVENTION

The present invention relates to a method of preparing a gene expression library enriched for genes encoding a class of proteins of interest comprising the steps of (i) preparing a cDNA expression library consisting of cDNA clones produced from mRNA enriched for sequences encoding the class of proteins; (ii) screening the cDNA expression library with a polyclonal antiserum directed toward the class of proteins; and (iii) pooling cDNA clones reactive with the antiserum.

For example, and not by way of limitation, the present invention provides for a method of preparing a gene expression library enriched for genes encoding secretory proteins comprising the steps of (i) preparing a cDNA expression library consisting of cDNA clones produced from mRNA collected from membrane-bound polysomes of a secretory tissue or cell type; (ii) screening the cDNA expression library with a polyclonal antiserum directed toward proteins secreted by the tissue; and (iii) pooling cDNA clones reactive with the antiserum. Analogous methods, using selection procedures directed not only at the underlying nucleic acid but also at a functional attribute of the proteins of interest, may be used to produce gene expression libraries enriched for a desired class of proteins. In nonlimiting embodiments of the invention, such libraries may be enriched for cell cycle associated genes, genes associated with a particular subcellular location, tissue specific genes, infection related genes, disease associated genes, and so forth.

The present invention further relates to a method of preparing cDNA or genomic libraries enriched for clones containing DNA sequences encoding secretory proteins or integral membrane proteins comprising the steps of (i) preparing a probe from mRNA collected from membrane-bound polysomes of a secretory tissue; (ii) screening the genomic or cDNA library with the probe of step (i); (iii) identifying clones that selectively hybridize with the probe; and (iv) pooling the clones that selectively hybridize with probe.

Following the identification of positive clones, the following method is used to eliminate false positive clones: (i) two probes are prepared, wherein one probe is prepared from mRNA collected from membrane-bound polysomes and the second probe

is prepared from mRNA collected from free polysomes; and (ii) screening the identified positive clones with the probe of step (i); wherein clones that selectively hybridize to the membrane-bound polysome mRNA derived probe and not the free polysome mRNA derived probe represent clones containing DNA sequences encoding either a secretory protein or an integral membrane protein. Following identification, the positive clones are pooled to form a library enriched for DNA sequences encoding secretory or integral membrane proteins.

4. DESCRIPTION OF THE FIGURES

FIGURE 1. Flow chart depicting a protocol according to the invention.

FIGURE 2. In GH3 rat pituitary cells, the mRNA's for prolactin (PRL), growth hormone (GH), and the epidermal growth factor receptor (EGFR) are localized virtually completely on membrane-bound polysomes. In contrast, the mRNA for actin are localized on free polysomes.

FIGURE 3. In LNCaP prostate cancer cells, prostate-specific antigen (PSA) mRNA is localized virtually completely on membrane-bound polysomes. The PSA mRNA analysis (top) represents a semi-quantitative assay for PSA mRNA levels, while the β-actin mRNA analysis is a non-quantitative assay for the integrity of input RNA. Lane 2 represents PSA mRNA detected in membrane bound polysome RNA, and lane 3 represents PSA mRNA levels present in free polysome RNA.

5. DETAILED DESCRIPTION OF THE INVENTION

By way of example, and not limitation, the present invention will first be described as it relates to the preparation of a gene expression library enriched for secretory proteins. Next, enriched libraries encoding other classes of proteins will be discussed. From this disclosure, the application of the method of the present invention to additional classes of proteins will become apparent.

When the present invention is applied to produce a gene expression library enriched for genes encoding proteins secreted by a particular secretory tissue, said

secretory tissue may be human or non-human in origin. Examples of secretory tissues include, but are not limited to, adipose tissue, thymus, bone, bone marrow (i.e., osteoblasts, osteoclasts, and cartilage tissue), placenta, breast, ovary, testes, stomach, intestine, pancreas, gall bladder, liver, pituitary, adrenal gland, salivary gland and other exocrine and endocrine tissues. Cells which secrete a protein of interest are also included in the definition of secretory tissue as applied herein; for example, lymphocytes, hematopoietic stem cells, endothelial cells, platelets and cells of the nervous system which produce neurotrophic factors, endorphins and/or neurotransmitters are considered to be secretory tissue according to the invention. The use of the term "tissue" does not necessarily imply that the cells form a contiguous multicellular structure. in addition, cell lines may be used according to the invention. In a specific, nonlimiting embodiment, the human breast cancer line MDA-MB-231, as described in el-Yazidi et al., 1995, Anticancer Res. 15:783-790, or similar cell lines, may be used.

The first phase of the method of the present invention comprises preparing a cDNA expression library consisting of cDNA clones produced from mRNA collected from membrane-bound polysomes of the secretory tissue. The secretory tissue may be derived directly from the subject or may be a cultured explant. Membrane-bound polysomes may be prepared using any method known in the art. Nonlimiting examples of such methods are set forth in Adesnik and Maschio, 1981, Eur. J. Biochem. .114:271-284 and Craig et al., 1979, Biochem. J. 181:737-756. Briefly, such method would comprise the harvest of secretory tissue, tissue homogenization (optionally in the presence of an RNase inhibitor), differential centrifugation on sucrose step gradients and collection of membrane-bound polysome fractions. The membrane-bound polysomes may then be subjected to a high salt wash to remove adventitiously bound free polysomes. (See also Collagen et al., eds., 1995, Current Protocols in Protein Science, John Wiley and Sons, Inc. New York, Unit 4). Since most of the mRNAs for intracellular proteins typically are present on free polysomes, while most of the mRNAs for secretory proteins are generally associated with membrane-bound polysomes, the fraction(s) prepared as described in the preceding paragraph may be significantly enriched for

mRNAs encoding secretory proteins.

Next, mRNA may be prepared from membrane-bound polysomes, and may be used, by standard techniques, to produce a cDNA expression library. Such a library may be constructed using various types of expression vectors, including but not limited to a bacteriophage λ-based expression vector such as λgt11 (Ausubel et al., eds., 1996, Current Protocols in Molecular Biology, John Wiley and Sons, Inc., New York) or Lambda ZAP® II Vector (Stratagene), a bacterial plasmid expression vector, a yeast shuttle expression vector, or a mammalian expression vector (see Ausubel, 1996, supra). In one nonlimiting embodiment, the present invention provides for a membrane-bound polysome generated gene expression library prepared using this first phase of the selection procedure and not the antibody-screening phase.

Where the second phase of the inventive method is applied, secreted proteins are prepared and then used to generate polyclonal antiserum or monoclonal antiserum. Such secreted proteins may be prepared from a directly obtainable biological product (for example, saliva, bile, gastrointestinal contents, etc.) or from a culture containing the secretory tissue (for example, a tissue explant culture). Where tissue explants are to be used, it should be noted that the presence of serum in the culture medium may result in severe contamination of the proteins secreted by the tissue explant, and, consequently, the generation of a highly non-specific antiserum. Therefore, it is preferred that a defined, serum-free medium be used for the incubation of explant cultures. The presence, in the incubation medium, of a limited number of defined proteins (for example, insulin in an adipose tissue explant culture) would not be expected to be problematic, both because the tissue under study may not contain mRNA encoding that defined protein and because any cDNA clones corresponding to defined proteins may be identified and removed from the product cDNA library and thereby eliminate from further consideration. Further, such a defined protein may serve as a carrier for the secreted proteins and may perform as a stabilizing agent and to inhibit their adherence to processing vessel walls. It may further be desirable to include a protease inhibitor in the explant culture, provided that care be taken that the presence of such inhibitor does not

substantially impair the metabolism cultured tissue.

The following examples relating to explant cultures, useful for generating secretory proteins to be used in antiserum production, may also be useful in the preparation of nucleic acid according to the first phase of the method of the invention.

In one particular nonlimiting embodiment of the invention, explant cultures of adipose tissue may be prepared from mammary adipose tissue obtained during surgical breast reduction (Petruschke et al., 1994, Int. J. Obes. Relat. Metab. Disord. 18:532-536; Hauner et al., 1989, J. Clin. Invest. 84:1663-1670) or from abdominal adipose tissue (Entemann and Hauner, 1996, Am. J. Physiol. 270 (Cell Phsyiol. 39):C1001-C1016). In related nonlimiting examples, adipocytes may be prepared according to methods described by Wabitsch et al., 1996, Metabol. 45:34-42 or MacDougald et al., 1995, Proc. Natl. Acad. Sci. U.S.A. 92:9034-9037.

In another nonlimiting embodiment of the invention, thymic explants may be prepared, for example as described in Meilin et al., 1992, Immunol. 77:208-213 and Schreier et al., 1991, Immunol. 24:621-629, wherein thymic epithelial cell explants obtained from children undergoing corrective cardiovascular surgery are cultured on extracellular matrix coated culture plates in a defined serum free medium supplemented with defined growth factors.

In yet another nonlimiting embodiment of the invention, bone marrow explant cultures may be prepared. For example, a description of the preparation of such cultures may be found in Bendall et al., 1994, Exp. Hematol. 22:1252-1260 which describes the isolation of bone marrow from acute myeloid leukemia patients followed by the preparation of two types of explant cultures containing either long-term bone marrow stroma or bone marrow fibroblasts. Evidence was provided by Bendall et al. which indicates that, upon incubation in a defined serum free medium supplemented with known cytokines, either type of explant culture may release factors capable of inhibiting leukemic cell death and/or in maintaining clonogenicity of malignant myeloid progenitors. In other examples, bone marrow explant cultures are described in Pribyl and LeBien, 1996, Proc. Natl. Acad. Sci. U.S.A. 93:10348-10353; Wolf et al., 1991, J.

Immunol. 147:3324-3330 and Dittel and LeBien, 1995, J. Immunol. 154:58-67, which relate to the isolation, from bone marrow obtained from aborted fetuses, of explant cultures of B cell precursors, hematopoietic stem cells, and bone marrow stromal cells, and the incubation of the cells in defined serum free medium. Gene expression libraries obtained from such cells may be useful toward isolating a gene encoding a hitherto unidentified soluble factor which acts in concert with interleukin-7 to promote the growth of CD10+/surface IgM- cells.

In a further nonlimiting embodiment of the invention, placental tissue explant cultures may be prepared as described in Mauro et al., 1992, J. Clin. Endocrinol. Metab. 75:1362-1367; Mauro et al., 1995, Reprod. Fertil. Dev. 7:1465-1470 and Larrea et al., 1993, J. Steroid Biochem. Mol. Biol. 46:497-505, which relate to the preparation of placental explants from trophoblastic tissues obtained from human placentas and incubation of the cells in a defined serum free medium supplemented with known growth factors.

Secreted fluid or culture medium from explant cultures may then be collected in the presence of appropriate protease inhibitors (such as serine protease inhibitors (e.g., PMSF), thiol protease inhibitors (e.g., leupeptin) and/or acidic protease inhibitors (e.g., pepstatin)) and subjected to centrifugation (for example, at 60,000 x g for one hour to yield an "S-100" supernatant) , and thus remove any tissue and/or cellular debris such as membranes released by any degraded tissue and/or cells. The population of secreted protein may then be concentrated using any standard procedure (e.g., filtration through an Amicon filter, or acetone precipitation followed by centrifugation). If the defined serum-free medium employed for a particular application contains a sufficiently high concentration of a particular defined protein(s), the presence of such a protein(s) may interfere with either the above concentration step and/or the production of antibodies against the proteins secreted by the tissue employed. In that case, it may prove advantageous to remove one or more of these defined protein(s) prior to the concentration step, employing a technique for specifically removing such a protein. This may be done a number of ways, including, for example, use of an affinity procedure employing either

immobilized antibodies specific for that protein, or a chemical reagent displaying a degree of specificity for the defined protein(s). An example of such a specific chemical reagent that may prove useful is blue Sepharose (Pharmacia Biotech), which has been employed for purification of interferon from a serum-free medium (Knight and Fahey, 1981, J. Biol. Chem. 256:3609-3611).

Standard immunological techniques, such as those described in Harlow and Lane (1988, *Antibodies: A Laboratorv Manual*, Cold Spring Harbor Laboratory, New York) may then be employed to prepare antiserum to the secreted proteins. For example, but not by way of limitation, it may be desirable, in order to improve the immune response in an inoculated animal to the secreted proteins, to couple the secreted proteins to an appropriate carrier protein (as described, for example, at p.77 of Harlow and Lane, supra) and/or to effect limited protease digestion of the secreted proteins. Next, one or more standard laboratory animals may be immunized with secreted proteins and antiserum may be recovered (see, for example, Chapter 5 of Harlow and Lane, supra). The success of antibody preparation may be evaluated by comparing the patterns of (i) immunoblots of the secreted protein population obtained using the antiserum as a probe (see Chapter 12 of Harlow and Lane, supra) with (ii) the pattern observed when the secreted protein population is subjected to gel electrophoresis under the same conditions used in (i), and then visualized by a detection technique such as silver staining.

Alternatively, antibodies against the secreted protein fraction may be prepared by phage display technology (see, for example,Winter et al., 1994, Annu. Rev. Immunol. 12:433-455; Marks et al.,1993, Bio/Technol. 11:1145-1149). Such technology permits the selection, from an antibody fragment gene recombinant bacteriophage library prepared from natural or synthetic antibody gene repertoires, of bacteriophage expressing an antibody fragment against a particular antigen. As a specific, nonlimiting example, the secreted protein fraction may be used to screen an antibody fragment gene recombinant bacteriophage library. Clones identified could then be recovered. False positives could be reduced or eliminated, for example, by preadsorbing the library with a preparation of intracellular or integral membrane proteins (i.e., non-secretory proteins) or by retesting

recovered clones for reactivity with non-secretory proteins, wherein such reactive clones are discarded. Antibody fragments produced by the bacteriophage clones could then be expressed in bacteria and recovered and then employed, either individually or as a pool of antibody fragments, for the screening of the cDNA expression library.

It may be noted that it is possible that an antiserum may not contain antibodies against all the secreted proteins. However, the success in raising antibodies against at least some of the proteins secreted by the tissue may be examined both by the procedure described above and/or by testing the ability of the antiserum to detect proteins known to be secreted by the tissue. If preliminary tests of the antiserum indicate that contamination by antibodies to intracellular and/or integral membrane proteins is a problem, standard immunological adsorption techniques may be employed to remove or reduce such contaminants. This may involve, for example, adsorption of the antiserum with an acetone powder prepared from intracellular and/or integral membrane proteins, such as is described on pages 632-633 of Harlow and Lane (supra).

Next, the cDNA expression library prepared according to the first phase of the invention may be screened with the antiserum directed toward proteins secreted by the tissue. There presently exist standard procedures for employing antiserum directed against a single protein to screen a cDNA library for a clone(s) corresponding to that protein (for example, the procedure for screening a bacteriophage cDNA expression library described in Unit 6.7 of-Ausubel et al., eds., 1996, Current Protocols in Molecular Biology, John Wiley and Sons, Inc., New York). Such methods may analogously be applied to identify clones which react with the polyclonal antiserum, directed against the population of secreted proteins. Recombinant cDNA clones observed to react with the antiserum, but not with preimmune serum from the same animal(s), may be selected for further use, and preferably may be retested to ascertain selective reactivity. Clones exhibiting a positive reaction may then be pooled, and then collectively represent a gene expression library enriched for genes encoding secretory proteins.

In a further nonlimiting embodiment, a protein expressed by one of the clones reactive with the polyclonal antiserum may then be used to select, from the

antiserum, antibodies specifically directed toward that protein, by a modification of the standard procedure as described on pages 312-315 of Harlow and Lane (supra) for immunoaffinity purification of antibodies on an antigen column.

In yet other embodiments of the invention, polyclonal antiserum prepared as set forth above may be used to screen a standard cDNA library (e.g., one which has not been enriched for sequences encoding a specific class of proteins, as set forth above). Reactive clones may then be pooled to create a single-tier enriched library. Such methods may be appropriate where enrichment at the nucleic acid level is problematic.

It should also be noted that an antiserum may be designed, according to the invention, to react with clones expressing proteins that are to be selected against, and then used to produce a library expressing proteins which are not recognized by the antiserum.

In yet another embodiment of the invention, in the first phase of the inventive method, a standard genomic or cDNA library is initially screened with a probe prepared from mRNA collected from membrane-bound polysomes of secretory tissue from which the library was generated. To eliminate false positives, the positive clones identified in the first phase are subjected to a second screening procedure. In the second phase of the inventive method, the positive clones are screened a second time with probes prepared either from membrane-bound polysome mRNA or free polysome mRNA prepared according to standard techniques, e.g., RT-PCR. Clones containing DNA fragments encoding secretory proteins or integral membrane proteins will differentially hybridize with the probes prepared from membrane-bound mRNA but not with probes prepared from free polysome mRNA. Such clones are pooled to yield a library enriched for genes encoding secretory proteins or integral membrane proteins.

When the present invention is applied to the screening of cDNA libraries, standard pre-existing libraries prepared from total cellular mRNA isolated from a specific type of tissue or cultured cell line, of the type purchased from, e.g., Stratagene, or readily prepared according to standard techniques, can be used. In addition, when screening genomic libraries standard genomic libraries can be generated using standard methods

well known to those skilled in the art. For a review of cloning strategies see, e.g., Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, (Volumes 1-3) Cold Spring Harbor Press; and Ausubel et al., 1989, Current Protcols in Molecular Biology, Green Publishing Associates and Wiley Interscience, N.Y.

Where the inventive method involves the differential screening of a library using a probe prepared from purified membrane-bound polysome mRNA, or free polysome mRNA, the probe can be prepared using a variety of different methods. Membrane-bound polysomes and free polysomes may be first prepared using techniques well known to those of skill in the art, such as those described in Craig et al., 1979, Biochem. J. 181:737-756. RNA may be prepared from each of the purified fractions using standard methods for purification of RNA from a sample.

A number of different methods are available for producing labeled hybridization probes from the purified RNA samples. One such method includes the PCR amplification method described by Brady et al., 1990, Methods in Molecular and Cellular Biology 2:17-25 for amplification of cDNA corresponding to the purified RNA. Probes can be prepared, for example, by carrying out a PCR reaction in the presence of radio-labeled nucleotides.

The labeled probes can be used to screen a cDNA library constructed from mRNA obtained from a cell type or tissue of interest. Alternatively the labeled fragment can be used to screen a genomic library derived from a cell of interest. Hybridization conditions should be sufficiently stringent as to eliminate any non-specific binding. For guidance regarding such conditions see, for example, Sambrook, et al., 1989, Molecular Cloning, A Laboratory Manual, Second Edition, Cold Spring Harbor Press, N.Y.; and Ausubel, et al., 1989, Current Protocols in Molecular Biology, Green Publishing Associates and Wiley Interscience, N.Y. Dot- or slot- blotting methods, such as those described in Unit 2.9B of Short Protocols in Molecular Biology, 3rd edition, 1995, Ausubel et al., Eds John Wiley and Sons, can also be used to perform differential hybridization screening assays to identify clones encoding secretory protein, *i.e.*, clones that hybridize to a probe prepared from membrane-bound polysome mRNA but not free

polysome mRNA.

In yet another embodiment of the invention, the DNA probe to be used in screening of libraries is labeled in the presence of fluorescent-labeled nucleotides of the type used in automated sequencing (see, e.g., Metzker, et al., 1996, Science 271:1420-1422). When utilizing such fluorescent-labeled nucleotides, each type of probe is labeled with a different color. For example, the probes prepared from, respectively, membrane-bound polysome mRNA or free polysome mRNA can be prepared employing fluorescent-labeled nucleotides that are, respectively red and green.

When screening a library using fluorescent-labeled probes, the library to be screened can be prepared on a support in an array. For example, the array could be prepared by transfer of a bacteriophage, cosmid, or plasmid library to a support appropriate for detecting hybridized colored probes, such as a nitrocellulose filter support (see, e.g., Short Protocols in Molecular Biology, 1995 Ausubel et al., eds. Wiley and Sons), or by a number of other techniques, such as the Gene Expression Mico-Array detection technique of Synteni. In using such an array, hybridization can be carried out simultaneously with both the green and red probes. Individual clones encoding secretory proteins or integral membrane proteins would be expected to fluoresce a green color. Clones encoding intracellular proteins would be expected to fluoresce a red color. Thus the use of fluorescent nucleotides for generation of labeled probes provides a direct method to identify genes encoding secretory proteins or integral membrane proteins.

From the foregoing nonlimiting example of how the present invention may be applied to obtain a gene expression library enriched for genes encoding secretory proteins, the variation of the inventive method for obtaining expression libraries enriched for genes encoding other specific classes of proteins would be readily apparent to the skilled artisan.

As a first nonlimiting example, a gene expression library may be prepared which is enriched for proteins associated with a particular stage of the cell cycle. mRNA may be prepared from a synchronized population of cells at a desired point in the cell cycle. Optionally, mRNAs representing constitutively expressed genes may be removed

by selective hybridization with mRNA (or corresponding cDNA) collected from cells at another stage of the cell cycle. Optionally, from the stage of the cell cycle of interest may be used to prepare a cDNA expression library. Further, proteins may be prepared from a synchronized population of cells at the same stage of the cell cycle, and used to generate an antiserum. It may be desirable to remove antibodies reactive with constitutive proteins by adsorption of the antiserum with proteins prepared from cells at one or more different stages of the cell cycle. The resulting antiserum may then be used to select clones from the cDNA library for incorporation into a library enriched for cell-cycle associated genes. A similar approach may be used to produce a library enriched for genes associated with apoptosis.

As a second nonlimiting example, the present invention may be used to prepare a gene expression library associated with a particular subcellular location. For example, mRNAs encoding proteins having nuclear localization signals may be identified by the presence of nucleic acid sequences corresponding to such signals, and then used to produce a cDNA expression library. Antiserum may be prepared toward nuclear proteins and then used to select cDNA clones which, taken together, constitute a library enriched for genes encoding nuclear proteins. Of note, it may be problematic to prepare a cDNA library enriched for clones encoding nuclear proteins. However, a polyclonal antiserum may be prepared against nuclear proteins and then used to screen a general cDNA expression library prepared using non-selected mRNAs.

As a third nonlimiting example, the present invention may be used to produce a gene expression library enriched in tissue specific proteins. mRNA may be prepared from the tissue type of interest; optionally, constitutively expressed mRNAs may be removed by selective hybridization to mRNA (or corresponding cDNA) from cells of a different type of tissue; and a cDNA expression library may be prepared. Proteins prepared from the tissue may then be used to raise an antiserum (which is optionally preadsorbed with proteins from a different tissue), and the antiserum may be used to select cDNA clones and thereby create a library enriched for tissue specific genes. An analogous method may be used to produce a library enriched for genes active during a

particular developmental stage.

As a fourth nonlimiting example, the present invention may be used to produce a gene expression library enriched for genes encoding proteins expressed by an infectious agent. Nucleic acid may be produced from the infectious agent itself or from a cell or organism infected with the infectious agent (optionally, and where appropriate, nucleic acids not associated with the infectious agent or infection may be removed by selective hybridization), and used to prepare a cDNA library. Proteins from the infectious agent or from a cell or organism infected with the infectious agent may then be harvested and used to raise an antiserum (which may optionally be preadsorbed with proteins not associated with the infectious agent or infection). The antiserum may then be used to select cDNA clones and thereby produce a library enriched ir genes encoding proteins expressed by the infectious agent oi associated with infection.

As a fifth nonlimiting example, the present invention may be used to prepare a gene expression library enriched for genes associated with a non-infectious disease mRNA may be prepared from a diseased cell, tissue, or organism (optionally mRNAs not associated with the disease may be removed by selective hybridization with mRNA (or corresponding cDNA) from a cell, tissue or organism which is not afflicted with the disease), and may be used to produce a cDNA expression library. Protein may be prepared from a diseased cell, tissue, or organism and used to raise an antiserum (which may optionally be preadsorbed with proteins prepared from a cell, tissue, or organism not afflicted with the disease) and the antiserum may be used to select clones which may be assembled to form a gene expression library enriched for disease-associated genes.

Similar methods may be used to prepare libraries enriched for genes associated with a particular developmental stage, aging, tissue and/or cellular damage or stress, genes induced by exposure of cells or tissues to hormones or other inducing agents, genes repressed by exposure of cells or tissues to hormones or other repressor agents, and genes induced (or repressed) by the expression of genes introduced by an infectious agent or by oncogenesis. Also, various cell types may be used according to the invention, including cells or cell lines of non-secretory tissues such as, for example, but

not by way of limitation, heart or striated muscle.

In yet another embodiment of the invention, the methods of the invention may be used to produce a library having a collection of cDNAs with more than one desirable characteristic. For example, a collection of cDNAs may be obtained that both encode secretory or membrane proteins and possess additional desirable properties, such as tissue specific expression, inducibility or repression by hormones or other agents, expression at specific stages of progression of cancer or development, or expression by an infectious agent, aging or stress. For example, but not by way of limitation, cDNA libraries can be obtained that are enriched for clones encoding secretory proteins using any of the methods described above. In addition, this step of the inventive method may be combined with preparation of RNA from tissue of interest and a different tissue as a source of material for preparation cDNA libraries. Such clones can then be used in differential hybridization protocols to identify clones that encode a tissue specific secretory or membrane bound protein.

6. EXAMPLE: ISOLATION OF DNA CLONES ENCODING SECRETORY OR MEMBRANE BOUND PROTEINS

The following section describes the isolation of mRNA molecules encoding secretory or integral membrane proteins using a differential hybridization protocol.

6.1. CELL FRACTIONATION, PREPARATION OF RNA AND PREPARATION OF PROBES

Membrane-bound polysomes were prepared from GH_3 rat pituitary cells using the "Preparation of high salt membrane-bound polyribosomes" technique described in Craig et al., 1979, Biochem. J., 181:737-756). Free polyribosomes were prepared from the GH_3 rat pituitary cells as described in Craig et al., (1979, *supra*). RNA was prepared from each of the cellular fractions as also described in Craig et al., (1979, *supra*), except

that Prime RNase inhibitor from 5'-3' , Inc. was added to inhibit RNAse activity. The PCR amplification method described by Brady et al., (1990, Methods in Molecular and Cellular Biology 2:17-25) was used for production of cDNA from either membrane-bound polysome or free polysome mRNA. RNA from either fraction was employed for synthesis of first strand cDNA, poly(dA) addition, and PCR amplification as described by Brady et al., with the exception that total RNA rather than poly(A)- containing RNA was employed as template. In addition the primer sequence employed by Brady et al. for X in the oligo(dT)-X primer was replaced with the following sequence: CAT GTC GTC CAG GCC GCT CTG GGA CAA AAT ATG AAT TC, which gave a greater yield of PCR product than the X sequence employed by Brady.

^{32}P-labeled probes were prepared by further PCR amplification of the cDNA products as follows: each PCR product was subjected to PCR (one cycle of 5 min/80° C, 5 min/95° C; 7 cycles of 1 min/94° C, 1.5 min/42° C) in the presence of either [^{32}P]dCTP plus [^{32}P]dATP in the primary screening, or [^{32}P]dCTP alone in the secondary and tertiary screening.

6.1.2. SCREENING OF LIBRARIES

A commercially available Stratagene GC rat pituitary cell lambda ZAPII cDNA library was titred, plated, and screened by hybridization as described in the Stratagene Lambda ZAPII Library Instruction Manual (March 8, 1994). The hybridization probe employed was the ^{32}P-labeled cDNA prepared from membrane-bound polysome RNA as described above. Plaques that were positive on two successive nitrocellulose filter lifts were subjected to secondary and tertiary screening with the ^{32}P-labeled cDNA prepared vs membrane-bound polysome RNA as described in the above referenced Stratagene Instruction Manuel. Plaques that were positive following a tertiary screen were subjected to *in vivo* excision of pBluescript phagemids and the resultant plasmids were then propagated in bacteria, and purified therefrom according to standard techniques.

To confirm that the selected plasmids encode secretory or integral membrane proteins, a slot-blotting protocol as described in Unit 2.9B of Short Protocols in Molecular Biology", 3rd Edition , 1995, Ausubel et al., Eds. John Wiley and Sons, was performed. Each positive plasmid was slot blotted onto duplicate nitrocellulose filters, and the nitrocellulose filters treated as described in Unit 2.9B with the following modifications: denaturation and neutralization were done by submerging the filters in the appropriate solutions. Denaturation was carried out for 2 min in a solution of 1.5M NaCl/0.5 M NaOH. Neutralization was carried out for 5 min at room temp in 1.5 M NaCl/0.5 M Tris-HCl (pH 8.0). Following neutralization, the filters were rinsed by incubating for 30 seconds at room temp in the following rinse solution: 0.2 M Tris-HCl (pH 7.5), 2XSSC.

Each filter was then hybridized as described in unit 2.10 of Ausubel et al., (*supra*) with a ^{32}P-labeled probe prepared from either membrane-bound polysome RNA or free polysome RNA, prepared as described above, and employing ^{32}PdCTP as the source of radioactivity. Briefly, hybridization was carried out in an aqueous prehybridization/hybridization solution (5X SSC, 5X Denhardt solution, 1% (w/v) SDS, 100µg/ml denatured salmon sperm DNA) at 68°C overnight. The nitrocellulose membranes were washed as described in Unit 2.10 of Ausubel et al. (*supra*), *i.e.*, in 0.2XSSC/0.1% SDS at room temperature. In some instances, were high background signal was observed, the filters were also washed at 42°C in 0.2XSSC/0.1%SDS and if needed at 68°C in 0.1XSSC/0.1%SDS. Following washing the membranes where exposed to X-ray film.

7. EXAMPLE: RNA MOLECULES ENCODING SECRETORY OR INTEGRAL MEMBRANE PROTEINS ARE LOCALIZED ON MEMBRANE BOUND POLYSOMES

The following section describes experimental results demonstrating that mRNA molecules encoding secretory or integral membrane bound proteins are localized

to membrane-bound polysomes thereby supporting the methods of the present invention.

First, cDNA hybridization probes corresponding to membrane-bound polysome RNA and free polysome mRNA derived from GH3 rat pituitary cells were prepared as described above. The probes were utilized to analyze four known cDNAs corresponding to the secretory proteins, prolactin (PRL) and growth hormone (GH), the cell surface integral membrane protein epidermal growth factor receptor (EGFR) and the control intracellular actin protein.

Plasmids containing cDNAs for actin, EGFR, PRL or GH were denatured, slot-blotted onto duplicate nitrocellulose filters (2 μg of each plasmid per slot), air dried; and then denatured, neutralized, rinsed, and baked all as described in 2.9B of Ausubel et al. (1995, *supra*). Labeled probes corresponding to the mRNA in membrane-bound polysomes or free polysomes were prepared from GH3 rat pituitary cells as described in Section 6.1, *supra*., employing [^{32}P]dCTP as a source of label. Following separation of labeled cDNA from free [^{32}P]dCTP on a G50 Sephadex column, each of the duplicate filters was hybridized with equal counts per minute of labeled cDNA prepared from either membrane-bound polysomes ("Mem-bd polys") or free polysomes ("free polys"), employing the procedures described in Unit 2.10 of Ausubel et al., (1995) for prehybridization, hybridization and washing. The filters were then exposed to X-ray film for two hours at -80°C.

Following hybridization, it was observed that the cDNAs for actin exhibited hybridization that was virtually completely specific for the free polysome mRNA probe, while the cDNAs for PRL, GH, and EGFR each hybridized very strongly to the membrane-bound polysome mRNA probe, and showed undetectable hybridization to the free polysome mRNA probe. (Figure 2) The results demonstrate that mRNAs encoding secretory or cell surface proteins is present specifically in membrane bound polysome RNA.

In addition, membrane and post-membrane fractions where prepared from human prostate cancer cells and each fraction was analyzed to determine the location of mRNA encoding the well known secretory prostate protein, prostate specific antigen

(PSA). LNCaP prostate cancer cells were incubated with androgen as described (Robbins et al., 1996, The Prostate 29:362) and membrane fraction and post-membrane fractions prepared (Bancroft, 1973, Exp. Cell Res. 79:275). Detached cells swollen in hypotonic buffer were disrupted by Dounce homogenization, followed by a slow speed spin to remove nuclei, and then high speed centrifugation to yield a pellet and supernatant, representing, respectively, the membrane and post-membrane fractions. RNA was then prepared from each fraction as described in Bancroft et al., (1973, Proc. Natl. Acad. Sci USA 70:3646-3649) except that the solution employed to suspend the membrane fraction was 1 mM EDTA, .05% SDS. RNA prepared from each fraction was then analyzed for PSA mRNA by comparative polymerase chain reaction (PCR), as described (Miniati et al., 1996, Cancer Letters 104:137). Equal quantities (2μg) of total RNA from either fraction were subjected to 35 rounds of PCR, either with PSA-specific primers in the presence of a PSA competitor or with control β-actin specific primers , and then subjected to gel electrophoresis in the presence of HaeII-cleaved ϕX174 DNA markers (Figure 3, lane 1), followed by EtBr-staining of the gel. The PSA mRNA analysis represents a semi-quantitative assay for PSA mRNA levels, while the β-actin mRNA analysis is a non-quantitative assay for the integrity of input RNA (Miniati et al., *supra*). The results from these experiments are presented in Figure 3. As indicated, a strong band corresponding to PSA mRNA is detected in membrane-bound polysome RNA ("Memb-bd polys", lane 2) but such a band was virtually undetectable in free polysome RNA ("Free polys", lane 3). The results indicate that PSA mRNA is located virtually exclusively in the membrane fraction (i.e., on membrane bound polysomes) of the LNCaP cells.

Other applications of the present invention would be readily apparent to the skilled artisan. Various references are cited herein, the contents of which are hereby incorporated by reference in their entireties.

CLAIMS

1. A method of preparing a gene expression library enriched for genes encoding a class of proteins of interest comprising the steps of (i) preparing a cDNA expression library consisting of cDNA clones produced from mRNA enriched for sequences encoding the class of proteins; (ii) screening the cDNA expression library with a polyclonal antiserum directed toward the class of proteins; and (iii) pooling cDNA clones reactive with the antiserum.

2. The method according to claim 1, wherein the proteins of interest are cell cycle associated proteins.

3. The method according to claim 1, wherein the proteins of interest are proteins localized in a particular subcellular location.

4. The method according to claim 1, wherein the proteins of interest are tissue specific proteins.

5. The method according to claim 1, wherein the proteins of interest are proteins expressed by an infectious agent.

6. The method according to claim 1, wherein the proteins of interest are associated with a specific non-infectious disease.

7. A method of preparing a gene expression library enriched for genes encoding a class of proteins of interest comprising the steps of (i) preparing a cDNA expression library consisting of cDNA clones

produced from mRNA enriched for sequences encoding the class of proteins; (ii) screening the cDNA expression library with a polyclonal antiserum directed toward the class of proteins; and (iii) pooling cDNA clones reactive with the antiserum.

8. The method according to claim 1, wherein the proteins of interest are cell cycle associated proteins.

9. The method according to claim 1, wherein the proteins of interest are proteins localized in a particular subcellular location.

10. The method according to claim 1, wherein the proteins of interest are tissue specific proteins.

11. The method according to claim 1, wherein the proteins of interest are proteins expressed by an infectious agent.

12. The method according to claim 1, wherein the proteins of interest are associated with a specific non-infectious disease.

13. A method of preparing a gene expression library enriched for genes encoding secretory proteins comprising the steps of (i) preparing a cDNA expression library consisting of cDNA clones produced from mRNA collected from membrane-bound polysomes of a secretory tissue; (ii) screening the cDNA expression library with a polyclonal antiserum directed toward proteins secreted by the tissue; and (iii) pooling cDNA clones reactive with the antiserum.

14. The method according to claim 7 wherein the secretory tissue is adipose tissue.

15. The method according to claim 7 wherein the secretory tissue is endocrine or paracrine tissue.

16. The method according to claim 7 wherein the secretory tissue is hematopoetic tissue.

17. A gene expression library prepared by a method comprising the steps of (i) preparing a cDNA expression library consisting of cDNA clones produced from mRNA enriched for sequences encoding a class of proteins; (ii) screening the cDNA expression library with a polyclonal antiserum directed toward the class of proteins; and (iii) pooling cDNA clones reactive with the antiserum.

18. The library according to claim 11, wherein the proteins of interest are cell cycle associated proteins.

19. The library according to claim 11, wherein the proteins of interest are proteins localized in a particular subcellular location.

20. The library according to claim 11, wherein the proteins of interest are tissue specific proteins.

21. The library according to claim 11, wherein the proteins of interest are proteins expressed by an infectious agent.

22. The library according to claim 11, wherein the proteins of interest are associated with a specific non-infectious disease.

23. A gene expression library enriched for genes encoding secretory proteins prepared by a method comprising the steps of (i) preparing a cDNA expression library consisting of cDNA clones produced from mRNA collected from membrane-bound polysomes of a secretory tissue; (ii) screening the cDNA expression library with a polyclonal antiserum directed toward proteins secreted by the tissue; and (iii) pooling cDNA clones reactive with the antiserum.

24. The library according to claim 17, wherein the secretory tissue is adipose tissue.

25. The library according to claim 17, wherein the secretory tissue is endocrine or paracrine tissue.

26. The library according to claim 17, wherein the secretory tissue is hematopoetic tissue.

27. A method of preparing cDNA or genomic libraries enriched for clones containing DNA sequences encoding secretory proteins, or integral membrane proteins, comprising the steps of (i) preparing a probe from mRNA collected from membrane-bound polysomes of a secretory tissue fraction; (ii) screening a standard genomic or cDNA library with the probe of step (i); (iii) identifying clones that selectively hybridize with the probe; and (iv) pooling the clones that selectively hybridize with probe.

28. The method according to claim 27 wherein the secretory tissue is adipose tissue.

29. The method according to claim 27 wherein the secretory tissue is endocrine or paracrine tissue.

30. The method according to claim 27 wherein the secretory tissue is hematopoetic tissue.

31. A method for identifying cDNA or genomic clones encoding secretory proteins, or integral membrane proteins, comprising : (i) preparing two probes, wherein one probe is prepared from mRNA collected from membrane-bound polysomes and the second probe is prepared from mRNA collected from free polysomes; (ii) screening cDNA or genomic clones with the probes of step (i); wherein clones that selectively hybridize to the membrane-bound polysome mRNA derived probe and not the free polysome mRNA derived probe represent clones containing DNA sequences encoding a secretory protein or an integral membrane protein; and (iii) pooling the clones that selectively hybridize.

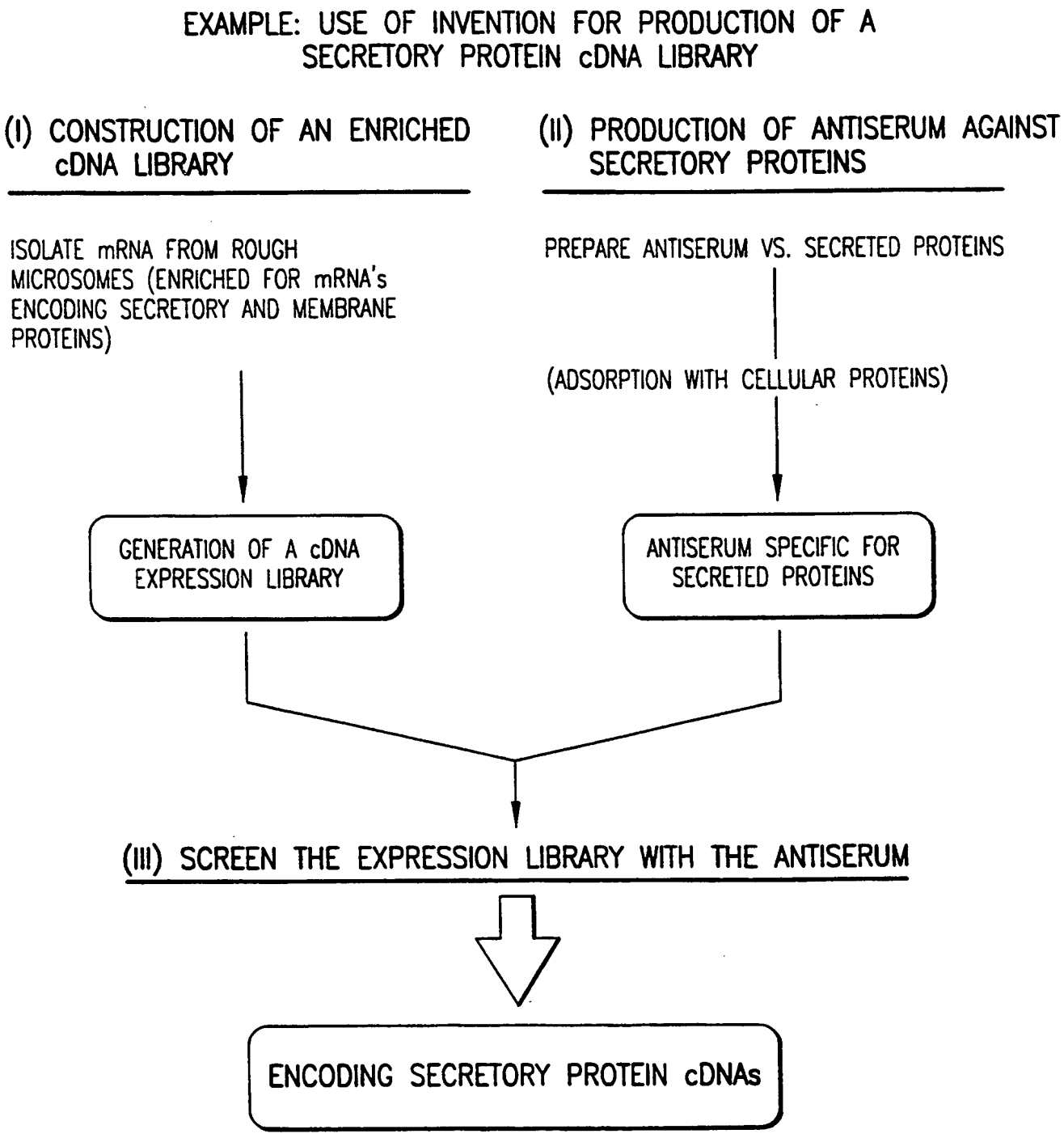

FIG.1

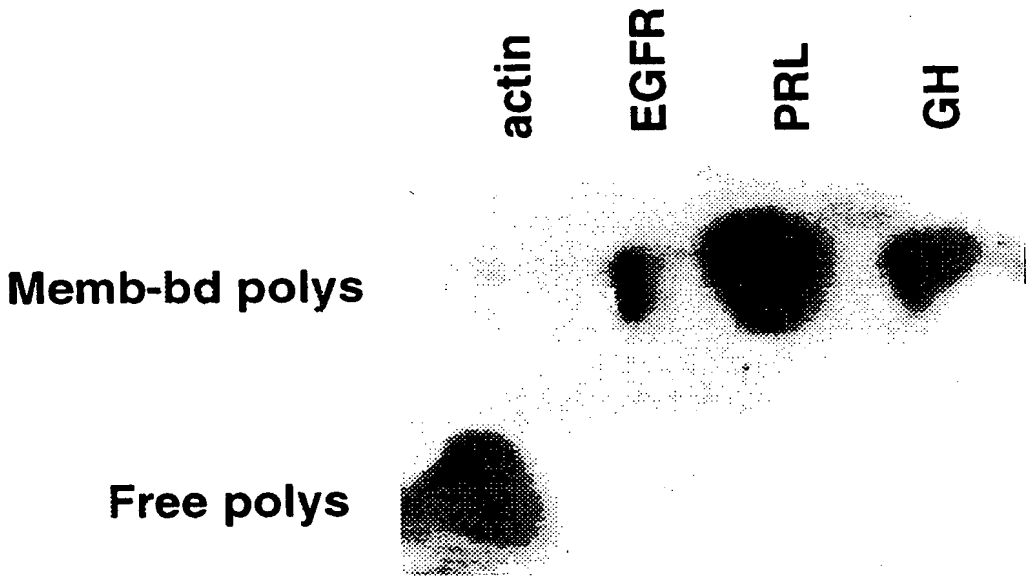

FIG.2

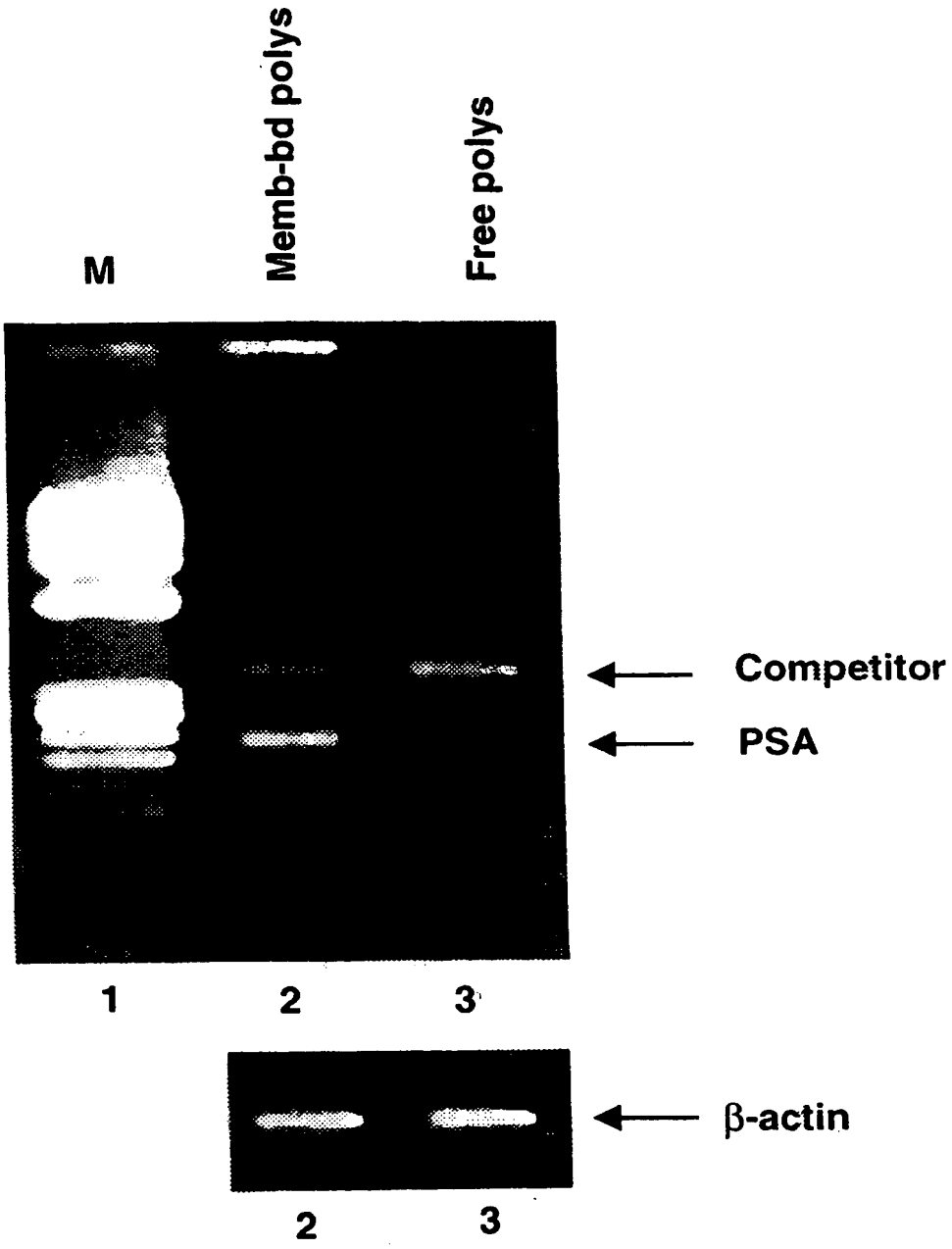

FIG.3

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/14759

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :C12N 15/63; C12P 19/34; C12Q 1/68, 1/70
US CL :435/5, 6, 320.1, 91.51

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/5, 6, 320.1, 91.51

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5,578,453 A (MCDONALD et al) 26 November 1996, columns 5-7.	1-31
Y	HEDRICK et al. Isolation of cDNA clones encoding T cell-specific membrane-associated proteins. Nature. 08 March 1984. Vol. 308. pages 149-153, see entire document, especially, page 149, column 2.	1-31
Y	KRAMER et al. Selective Cloning of cDNA for Secretory Proteins of Early Embryos. The Journal of Biological Chemistry. 11 March 1994. Vol. 269. No. 10. pages 7255-7261, see entire document, especially page 7256, column 1.	1-31

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance, the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance, the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 29 SEPTEMBER 1998	Date of mailing of the international search report 23 OCT 1998
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer TERRY A. MCKELVEY Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/14759

B. FIELDS SEARCHED
Electronic data bases consulted (Name of data base and where practicable terms used):

APS, DIALOG (biotechnology OneSearch databases)
search terms: subtractive hybridization, polyclonal, antisera, antiserum, antibod?, cDNA, librar?, screen?, rough microsome?, polysome?, select?, enrich?, secret?, express?, cell cycle, subcellular, tissue, specific, infect?, disease, pool?, adipose, endocrine, paracrine, hematopoietic